

0570
07/17

10



OIPE

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/941,193A

DATE: 07/23/2002 8-6
TIME: 13:04:51

Input Set : A:\Seqsub2.app

Output Set: N:\CRF3\07232002\I941193A.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: BROW, MARY ANN D.

7 LYAMICHEV, VICTOR I.

8 OLIVE, DAVID M.

10 (ii) TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
11 PATHOGENS

13 (iii) NUMBER OF SEQUENCES: 165

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: MEDLEN & CARROLL

17 (B) STREET: 220 MONTGOMERY STREET, SUITE 2200

18 (C) CITY: SAN FRANCISCO

19 (D) STATE: CALIFORNIA

20 (E) COUNTRY: UNITED STATES OF AMERICA

21 (F) ZIP: 94104

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Floppy disk

25 (B) COMPUTER: IBM PC compatible

26 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

27 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

29 (vi) CURRENT APPLICATION DATA:

C--> 30 (A) APPLICATION NUMBER: US/09/941,193A

C--> 31 (B) FILING DATE: 28-Aug-2001

32 (C) CLASSIFICATION:

34 (viii) ATTORNEY/AGENT INFORMATION:

35 (A) NAME: CARROLL, PETER G.

36 (B) REGISTRATION NUMBER: 32,837

37 (C) REFERENCE/DOCKET NUMBER: FORS-01756

39 (ix) TELECOMMUNICATION INFORMATION:

40 (A) TELEPHONE: (415) 705-8410

41 (B) TELEFAX: (415) 397-8338

44 (2) INFORMATION FOR SEQ ID NO: 1:

46 (i) SEQUENCE CHARACTERISTICS:

47 (A) LENGTH: 2506 base pairs

48 (B) TYPE: nucleic acid

49 (C) STRANDEDNESS: double

50 (D) TOPOLOGY: linear

52 (ii) MOLECULE TYPE: DNA (genomic)

56 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

58	ATGAGGGGGA TGCTGCCCT CTTTGAGCC AAGGGCCGGG TCCTCCTGGT GGACGGCCAC	60
60	CACCTGGCCT ACCGCACCTT CCACGCCCTG AAGGGCCTCA CCACCAGCCG GGGGGAGCCG	120
62	GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCCTCAAGGA GGACGGGGAC	180
64	GCGGTGATCG TGGTCTTTGA CGCCAAGGCC CCCTCCTTCC GCCACGAGGC CTACGGGGGG	240

RAW SEQUENCE LISTING

DATE: 07/23/2002

PATENT APPLICATION: US/09/941,193A

TIME: 13:04:51

Input Set : A:\Seqsub2.app

Output Set: N:\CRF3\07232002\I941193A.raw

```

66 TACAAGGCGG GCCGGGCCCC CACGCCGGAG GACTTTCCCC GGCAACTCGC CCTCATCAAG      300
68 GAGCTGGTGG ACCTCCTGGG GCTGGCGCGC CTCGAGGTCC CGGGCTACGA GGCGGACGAC      360
70 GTCCTGGCCA GCCTGGCCAA GAAGGCGGAA AAGGAGGGCT ACGAGGTCCG CATCCTCACC      420
72 GCCGACAAAG ACCTTTACCA GCTCCTTTCC GACCGCATCC ACGTCCTCCA CCCCAGAGGG      480
74 TACCTCATCA CCCCGGCCTG GCTTTGGGAA AAGTACGGCC TGAGGCCCGA CCAGTGGGCC      540
76 GACTACCGGG CCCTGACCGG GGACGAGTCC GACAACCTTC CCGGGGTCAA GGGCATCGGG      600
78 GAGAAGACGG CGAGGAAGCT TCTGGAGGAG TGGGGGAGCC TGGAAGCCCT CCTCAAGAAC      660
80 CTGGACCGGC TGAAGCCCGC CATCCGGGAG AAGATCCTGG CCCACATGGA CGATCTGAAG      720
82 CTCTCCTGGG ACCTGGCCAA GGTGCGCACC GACCTGCCCC TGGAGGTGGA CTTGCGCAAA      780
84 AGGCGGGAGC CCGACCGGGA GAGGCTTAGG GCCTTTCTGG AGAGGCTTGA GTTTGGCAGC      840
86 CTCCTCCACG AGTTCGGCCT TCTGGAAAGC CCAAGGCCCT TGGAGGAGGC CCCCTGGCCC      900
88 CCGCCGGAAG GGGCCTTCGT GGGCTTTGTG CTTTCCCGCA AGGAGCCCAT GTGGGCCGAT      960
90 CTTCTGGCCC TGGCCGCCGC CAGGGGGGGC CGGGTCCACC GGGCCCCCGA GCCTTATAAA     1020
92 GCCCTCAGGG ACCTGAAGGA GGCGCGGGGG CTTCTCGCCA AAGACCTGAG CGTTCTGGCC     1080
94 CTGAGGGAAG GCCTTGGCCT CCCGCCGGGC GACGACCCCA TGCTCCTCGC CTACCTCCTG     1140
96 GACCCTTCCA ACACCACCCC CGAGGGGGTG GCCCGGCGCT ACGGCGGGGA GTGGACGGAG     1200
98 GAGGCGGGGG AGCGGGCCGC CTTTCCGAG AGGCTCTTCG CCAACCTGTG GGGGAGGCTT     1260
100 GAGGGGGAGG AGAGGCTCCT TTGGCTTTAC CGGGAGGTGG AGAGGCCCTT TTCCGCTGTC     1320
102 CTGGCCACAC TGGAGGCCAC GGGGGTGCGC CTGGACGTGG CCTATCTCAG GGCCTTGTCC     1380
104 CTGGAGGTGG CCGAGGAGAT CGCCCGCCTC GAGGCCGAGG TCTTCCGCCT GGCCGGCCAC     1440
106 CCCTTCAACC TCAACTCCCG GGACCAGCTG GAAAGGGTCC TCTTTGACGA GCTAGGGCTT     1500
108 CCCGCCATCG GCAAGACGGA GAAGACCGGC AAGCGCTCCA CCAGCGCCGC CGTCTGGAG     1560
110 GCCCTCCGCG AGGCCACCCC CATCGTGGAG AAGATCCTGC AGTACCGGGA GCTCACCAG     1620
112 CTGAAGAGCA CCTACATTGA CCCCTTGCCG GACCTCATCC ACCCCAGGAC GGGCCGCCTC     1680
114 CACACCCGCT TCAACCAGAC GGCCACGGCC ACGGGCAGGC TAAGTAGCTC CGATCCCAAC     1740
116 CTCCAGAACA TCCCCGTCCG CACCCCGCTT GGGCAGAGGA TCCGCCGGGC CTTTCATCGCC     1800
118 GAGGAGGGGT GGCTATTGGT GGCCCTGGAC TATAGCCAGA TAGAGCTCAG GGTGCTGGCC     1860
120 CACCTCTCCG GCGACGAGAA CCTGATCCGG GTCTTCCAGG AGGGGCGGGA CATCCACACG     1920
122 GAGACCGCCA GCTGGATGTT CGGCGTCCCC CGGGAGGCCG TGGACCCCTT GATGCGCCGG     1980
124 GCGGCCAAGA CCATCAACTT CGGGGTCTCT TACGGCATGT CGGCCACCGC CCTCTCCAG     2040
126 GAGCTAGCCA TCCCTTACGA GGAGGCCAG GCCTTCATTG AGCGCTACTT TCAGAGCTTC     2100
128 CCAAGGTGCG GGGCCTGGAT TGAGAAGACC CTGGAGGAGG GCAGGAGGCG GGGGTACGTG     2160
130 GAGACCCTCT TCGGCCGCCG CCGCTACGTG CCAGACCTAG AGGCCCGGGT GAAGAGCGTG     2220
132 CGGGAGGCGG CCGAGCGCAT GGCCTTCAAC ATGCCCCTCC AGGGCACCGC CGCCGACCTC     2280
134 ATGAAGCTGG CTATGGTGAA GCTCTTCCCC AGGCTGGAGG AAATGGGGGC CAGGATGCTC     2340
136 CTTCAGGTCC ACGACGAGCT GGTCCCTCGAG GCCCAAAGAG AGAGGGCGGA GGCCGTGGCC     2400
138 CGGCTGGCCA AGGAGGTCAT GGAGGGGGTG TATCCCCTGG CCGTGCCCTT GGAGGTGGAG     2460
140 GTGGGGATAG GGGAGGACTG GCTCTCCGCC AAGGAGTGAT ACCACC      2506
142 (2) INFORMATION FOR SEQ ID NO: 2:
144     (i) SEQUENCE CHARACTERISTICS:
145         (A) LENGTH: 2496 base pairs
146         (B) TYPE: nucleic acid
147         (C) STRANDEDNESS: double
148         (D) TOPOLOGY: linear
150     (ii) MOLECULE TYPE: DNA (genomic)
154     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
156 ATGGCGATGC TTCCCCTCTT TGAGCCCAAA GGCCGCGTGC TCCTGGTGA CGGCCACCAC      60
158 CTGGCCTACC GCACCTTCTT TGCCCTCAAG GGCTTCACCA CCAGCCGCGG CGAACCCGTT     120
160 CAGGCGGTCT ACGGCTTCGC CAAAAGCCTC CTCAAGGCCC TGAAGGAGGA CGGGGACGTG     180

```

RAW SEQUENCE LISTING

DATE: 07/23/2002

PATENT APPLICATION: US/09/941,193A

TIME: 13:04:51

Input Set : A:\Seqsub2.app

Output Set: N:\CRF3\07232002\I941193A.raw

```

162 GTGGTGGTGG TCTTTGACGC CAAGGCCCCC TCCTTCCGCC ACGAGGCCTA CGAGGCCTAC      240
164 AAGGCGGGCC GGGCCCCCAC CCCGGAGGAC TTTCCCGGCC AGCTGGCCCT CATCAAGGAG      300
166 TTGGTGGACC TCCTAGGCCT TGTGCGGCTG GAGGTTCCCG GCTTTGAGGC GGACGACGTG      360
168 CTGGCCACCC TGGCCAAGCG GCGGAAAAAG GAGGGGTACG AGGTGCGCAT CCTCACTGCC      420
170 GACCGCGACC TCTACCAGCT CCTTTCGGAG CGCATCGCCA TCCTCCACCC TGAGGGGTAC      480
172 CTGATCACCC CGGCGTGGCT TTACGAGAAG TACGGCCTGC GCCCGGAGCA GTGGGTGGAC      540
174 TACCGGGCCC TGGCGGGGGA CCCCTCGGAT AACATCCCCG GGGTGAAGGG CATCGGGGAG      600
176 AAGACCGCCC AGAGGCTCAT CCGCGAGTGG GGGAGCCTGG AAAACCTCTT CCAGCACCTG      660
178 GACCAGGTGA AGCCCTCCTT GCGGGAGAAG CTCCAGGCGG GCATGGAGGC CCTGGCCCTT      720
180 TCCCGGAAGC TTTCCAGGT GCACACTGAC CTGCCCCTGG AGGTGGACTT CGGGAGGCGC      780
182 CGCACACCCA ACCTGGAGGG TCTGCGGGCT TTTTGGAGC GGTGGAGTT TGGAAGCCTC      840
184 CTCCACGAGT TCGGCCTCCT GGAGGGGCGG AAGGCGGCAG AGGAGGCCCC CTGGCCCCCT      900
186 CCGGAAGGGG CTTTTTTGGG CTTTTCTTTT TCCCGTCCCG AGCCCATGTG GGCCGAGCTT      960
188 CTGGCCCTGG CTGGGGCGTG GGAGGGGCGC CTCCATCGGG CACAAGACCC CCTTAGGGGC     1020
190 CTGAGGGACC TTAAGGGGGT GCGGGGAATC CTGGCCAAGG ACCTGGCGGT TTTGGCCCTG     1080
192 CGGGAGGGCC TGGACCTCTT CCCAGAGGAC GACCCCATGC TCCTGGCCTA CCTTCTGGAC     1140
194 CCCTCCAACA CCACCCCTGA GGGGGTGGCC CGGCGTTACG GGGGGGAGTG GACGGAGGAT     1200
196 GCGGGGGAGA GGGCCCTCCT GGCCGAGCGC CTCTTCCAGA CCCTAAAGGA GCGCCTTAAG     1260
198 GGAGAAGAAC GCCTGCTTTG GCTTTACGAG GAGGTGGAGA AGCCGCTTTC CCGGTGTGTG     1320
200 GCCCGGATGG AGGCCACGGG GGTCCGGCTG GACGTGGCCT ACCTCCAGGC CCTCTCCCTG     1380
202 GAGGTGGAGG CGGAGGTGCG CCAGCTGGAG GAGGAGGTCT TCCGCTGGC CGGCCACCCC     1440
204 TTCAACCTCA ACTCCGCGA CCAGCTGGAG CGGGTGCTCT TTGACGAGCT GGGCCTGCCT     1500
206 GCCATCGGCA AGACGGAGAA GACGGGGAAG CGCTCCACCA GCGCTGCCGT GCTGGAGGCC     1560
208 CTCCGAGAGG ACCACCCCAT CGTGGACCGC ATCCTTCAGT ACCGGGAGCT CACCAAGCTC     1620
210 AAGAACACCT ACATAGACCC CTGCCCCGCC CTGGTCCACC CCAAGACCGG CCGGCTCCAC     1680
212 ACCCGCTTCA ACCAGACGGC CACCGCCACG GGCAGGCTTT CCAGCTCCGA CCCCACCTG     1740
214 CAGAACATCC CCGTGCGCAC CCCTCTGGGC CAGCGCATCC GCCGAGCCTT CGTGGCCGAG     1800
216 GAGGGCTGGG TGCTGGTGGT CTTGGAATAC AGCCAGATTG AGCTTCGGGT CCTGGCCAC     1860
218 CTCTCCGGGG ACGAGAACCT GATCCGGGTC TTTAGGAGG GGAGGGACAT CCACACCCAG     1920
220 ACCGCCAGCT GGATGTTCCG CGTTTCCCCC GAAGGGGTAG ACCCTCTGAT GCGCCGGGCG     1980
222 GCCAAGACCA TCAACTTCGG GGTGCTCTAC GGCATGTCCG CCCACCGCCT CTCCGGGGAG     2040
224 CTTTCCATCC CCTACGAGGA GCGGTGGCC TTCATTGAGC GCTACTTCCA GAGCTACCCC     2100
226 AAGGTGCGGG CCTGGATTGA GGGGACCCTC GAGGAGGGCC GCCGGCGGGG GTATGTGGAG     2160
228 ACCCTCTTCG GCCGCCGGCG CTATGTGCCC GACCTCAACG CCCGGGTGAA GAGCGTGCGC     2220
230 GAGGCGGCGG AGCGCATGGC CTTCAACATG CCGGTCCAGG GCACCGCCGC CGACCTCATG     2280
232 AAGCTGGCCA TGGTGCGGCT TTTCCCCCGG CTTCAGGAAC TGGGGGCGAG GATGCTTTTG     2340
234 CAGGTGCACG ACGAGCTGGT CCTCGAGGCC CCCAAGGACC GGGCGGAGAG GGTAGCCGCT     2400
236 TTGGCCAAGG AGGTCATGGA GGGGGTCTGG CCCCTGCAGG TGCCCCTGGA GGTGGAGGTG     2460
238 GGCCTGGGGG AGGACTGGCT CTCCGCCAAG GAGTAG                                     2496

240 (2) INFORMATION FOR SEQ ID NO: 3:
242 (i) SEQUENCE CHARACTERISTICS:
243 (A) LENGTH: 2504 base pairs
244 (B) TYPE: nucleic acid
245 (C) STRANDEDNESS: double
246 (D) TOPOLOGY: linear
248 (ii) MOLECULE TYPE: DNA (genomic)
252 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
254 ATGGAGGCGA TGCTTCCGCT CTTTGAACCC AAAGGCCGGG TCCTCCTGGT GGACGGCCAC      60
256 CACCTGGCCT ACCGCACCTT CTTGCCCCTG AAGGGCCTCA CCACGAGCCG GGGCGAACCG      120

```

RAW SEQUENCE LISTING

DATE: 07/23/2002

PATENT APPLICATION: US/09/941,193A

TIME: 13:04:51

Input Set : A:\Seqsub2.app

Output Set: N:\CRF3\07232002\I941193A.raw

```

258 GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCCTGAAGGA GGACGGGTAC 180
260 AAGGCCGTCT TCGTGGTCTT TGACGCCAAG GCCCCCTCCT TCCGCCACGA GGCCTACGAG 240
262 GCCTACAAGG CGGGGAGGGC CCCGACCCCC GAGGACTTCC CCCGGCAGCT CGCCCTCATC 300
264 AAGGAGCTGG TGGACCTCCT GGGGTTTACC CGCCTCGAGG TCCCCGGCTA CGAGGCGGAC 360
266 GACGTTCTCG CCACCTTGGC CAAGAAGGCG GAAAAGGAGG GGTACGAGGT GCGCATCCTC 420
268 ACCGCCGACC GCGACCTCTA CCAACTCGTC TCCGACCGCG TCGCCGTCCT CCACCCCGAG 480
270 GGCCACCTCA TCACCCCGGA GTGGCTTTGG GAGAAGTACG GCCTCAGGCC GGAGCAGTGG 540
272 GTGGACTTCC GCGCCCTCGT GGGGGACCCC TCCGACAACC TCCCCGGGGT CAAGGGCATC 600
274 GGGGAGAAGA CCGCCCTCAA GCTCCTCAAG GAGTGGGGAA GCCTGGA AAA CCTCCTCAAG 660
276 AACCTGGACC GGGTAAAGCC AGAAAACGTC CGGGAGAAGA TCAAGGCCCA CCTGGAAGAC 720
278 CTCAGGCTCT CTTGGAGCT CTCCCGGGTG CGCACCAGCC TCCCCCTGGA GGTGGACCTC 780
280 GCCCAGGGGC GGGAGCCCGA CCGGGAGGGG CTTAGGGCCT TCCTGGAGAG GCTGGAGTTC 840
282 GGCAGCCTCC TCCACGAGTT CGGCCTCCTG GAGGCCCCCG CCCCCCTGGA GGAGGCCCCC 900
284 TGGCCCCCGC CGGAAGGGGC CTTCTGTGGC TTCGTCCTCT CCCGCCCGA GCCCATGTGG 960
286 GCGGAGCTTA AAGCCCTGGC CGCCTGCAGG GACGGCCGGG TGCACCGGGC AGCAGACCCC 1020
288 TTGGCGGGGC TAAAGGACCT CAAGGAGGTC CGGGGCCTCC TCGCCAAGGA CCTCGCCGTC 1080
290 TTGGCCTCGA GGGAGGGGCT AGACCTCGTG CCCGGGGACG ACCCCATGCT CCTCGCCTAC 1140
292 CTCCTGGACC CCTCCAACAC CACCCCCGAG GGGGTGGCGC GCGCTACGG GGGGGAGTGG 1200
294 ACGGAGGACG CCGCCCACCG GGCCCTCCTC TCGGAGAGGC TCCATCGGAA CCTCCTTAAG 1260
296 CGCCTCGAGG GGGAGGAGAA GCTCCTTTGG CTCTACCAG AGGTGGAAAA GCCCCCTCTC 1320
298 CGGGTCCTGG CCCACATGGA GGCCACCGGG GTACGGCTGG ACGTGGCCTA CTTTACAGCC 1380
300 CTTTCCCTGG AGCTTGCGGA GGAGATCCCG CGCCTCGAGG AGGAGGTCTT CCGCTTGCGC 1440
302 GGCCACCCCT TCAACCTCAA CTCCCGGGAC CAGCTGGAAA GGGTGCTCTT TGACGAGCTT 1500
304 AGGTTTCCCG CCTTGGGGAA GACGCAAAAG ACAGGCAAGC GCTCCACCAG CGCCGCGGTG 1560
306 CTGGAGGCCC TACGGGAGGC CCACCCCATC GTGGAGAAGA TCCTCCAGCA CCGGGAGCTC 1620
308 ACCAAGCTCA AGAACACCTA CGTGGACCCC CTCCCAAGCC TCGTCCACCC GAGGACGGGC 1680
310 CGCCTCCACA CCCGCTTCAA CCAGACGGCC ACGGCCACGG GGAGGCTTAG TAGCTCCGAC 1740
312 CCAACCTGC AGAACATCCC CGTCCGCACC CCCTTGGGCC AGAGGATCCG CCGGGCCTTC 1800
314 GTGGCCGAGG CGGGTTGGGC GTTGGTGGCC CTGGACTATA GCCAGATAGA GCTCCGCGTC 1860
316 CTCGCCCACC TCTCCGGGGA CGAAAACCTG ATCAGGGTCT TCCAGGAGGG GAAGGACATC 1920
318 CACACCCAGA CCGCAAGCTG GATGTTCCGG GTCCCCCGG AGGCCGTGGA CCCCCTGATG 1980
320 CGCCGGGCGG CCAAGACGGT GAACTTCGGC GTCCTCTACG GCATGTCCGC CCATAGGCTC 2040
322 TCCCAGGAGC TTGCCATCCC CTACGAGGAG GCGGTGGCCT TTATAGAGGC TACTTCCAAA 2100
324 GCTTCCCCAA GGTGCGGGCC TGGATAGAAA AGACCCTGGA GGAGGGGAGG AAGCGGGGCT 2160
326 ACGTGGAAAC CCTCTTCGGA AGAAGGCGCT ACGTGCCCGA CCTCAACGCC CGGGTGAAGA 2220
328 GCGTCAGGGA GGCCGCGGAG CGCATGGCCT TCAACATGCC CGTCCAGGGC ACCGCCGCCG 2280
330 ACCTCATGAA GCTCGCCATG GTGAAGCTCT TCCCCGCTC CCGGGAGATG GGGGCCCGCA 2340
332 TGCTCCTCCA GGTCCACGAC GAGCTCCTCC TGGAGGCCCC CCAAGCGCGG GCCGAGGAGG 2400
334 TGGCGGCTTT GGCCAAGGAG GCCATGGAGA AGGCCTATCC CCTCGCCGTG CCCCTGGAGG 2460
336 TGGAGGTGGG GATGGGGGAG GACTGGCTTT CCGCAAGGG TTAG 2504

```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 832 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu

RAW SEQUENCE LISTING

DATE: 07/23/2002

PATENT APPLICATION: US/09/941,193A

TIME: 13:04:51

Input Set : A:\Seqsub2.app

Output Set: N:\CRF3\07232002\I941193A.raw

```

353      1          5          10          15
355  Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly
356                20          25          30
358  Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
359                35          40          45
361  Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val
362                50          55          60
364  Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly
365                65          70          75          80
367  Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu
368                85          90          95
370  Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu
371                100          105          110
373  Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys
374                115          120          125
376  Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp
377                130          135          140
379  Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly
380                145          150          155          160
382  Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro
383                165          170          175
385  Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn
386                180          185          190
388  Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu
389                195          200          205
391  Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu
392                210          215          220
394  Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys
395                225          230          235          240
397  Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val
398                245          250          255
400  Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe
401                260          265          270
403  Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
404                275          280          285
406  Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
407                290          295          300
409  Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp
410                305          310          315          320
412  Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro
413                325          330          335
415  Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu
416                340          345          350
418  Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro
419                355          360          365
421  Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn
422                370          375          380
424  Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu
425                385          390          395          400

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/23/2002
PATENT APPLICATION: US/09/941,193A TIME: 13:04:52

Input Set : A:\Seqsub2.app
Output Set: N:\CRF3\07232002\I941193A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 4,5,181,182,190,366,617,628,685,714,722,738,784,1022,1029
Seq#:7; N Pos. 1038,1053,1098,1105,1206,1227,1244,1251,1252,1253,1350,1380
Seq#:7; N Pos. 1497,1530,1569,1572,1641,1653,1655,1770,1812,2319,2346,2396
Seq#:8; Xaa Pos.2,63,109,186,205,209,227,228,233,240,243,244,247,260,290
Seq#:8; Xaa Pos.329,336,340,368,414,417,418,431,551,605,773,794,798,823,833

VERIFICATION SUMMARY

DATE: 07/23/2002

PATENT APPLICATION: US/09/941,193A

TIME: 13:04:52

Input Set : A:\Seqsub2.app

Output Set: N:\CRF3\07232002\I941193A.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:1126 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0
L:1135 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:48
L:1144 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:96
L:1159 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:176
L:1162 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:192
L:1165 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:208
L:1168 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:224
L:1171 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:240
L:1174 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:256
L:1180 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:288
L:1186 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:320
L:1189 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:336
L:1192 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:352
L:1201 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:400
L:1204 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:416
L:1228 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:544
L:1237 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:592
L:1270 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:768
L:1273 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:784
L:1279 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:816
L:1282 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:832